



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/797,553A  
Source: ITW  
Date Processed by STIC: 10-18-04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.2 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - cPAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 06/05/04):  
U.S. Patent and Trademark Office, 220 20<sup>th</sup> Street S., Customer Window, Mail Stop Sequence, Crystal Plaza Two, Lobby, Room 1B03, Arlington, VA 22202

Revised 05/17/04

# Raw Sequence Listing Error Summary

## ERROR DETECTED

## SUGGESTED CORRECTION

SERIAL NUMBER: 10/797,553A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics  
    Wrapped Aminos      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2      Invalid Line Length      The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino  
    Numbering      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4      Non-ASCII      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length      Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>.-<223> section that some may be missing.
- 6      PatentIn 2.0  
    "bug"      A "bug" in PatentIn version 2.0 has caused the <220>.-<223> section to be missing from amino acid sequences(s)     . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>.-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>.-<223> sections for Artificial or Unknown sequences.
- 7      Skipped Sequences  
    (OLD RULES)      Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
    (2) INFORMATION FOR SEQ ID NO X (insert SEQ ID NO where "X" is shown)  
    (i) SEQUENCE CHARACTERISTICS (Do not insert any subheadings under this heading)  
    (xi) SEQUENCE DESCRIPTION SEQ ID NO X (insert SEQ ID NO where "X" is shown)  
    This sequence is intentionally skipped  
  
    Please also adjust the "(ii) NUMBER OF SEQUENCES" response to include the skipped sequences.
- 8      Skipped Sequences  
    (NEW RULES)      Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
    <210> sequence id number  
    <400> sequence id number  
    000
- 9      Use of n's or Xaa's  
    (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
    Per 1.823 of Sequence Rules, use of <220>.-<223> is MANDATORY if n's or Xaa's are present.  
    In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10      Invalid <213>  
    Response      Per 1.823 of Sequence Rules, the only valid <213> responses are Unknown, Artificial Sequence, or scientific name (Genus/species). <220>.-<223> section is required when <213> response is Unknown or is Artificial Sequence.
- 11      Use of <220>      Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
    (Sec. "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0  
    "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13      Misuse of n/Xaa      "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid.



IFWO

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/797,553A

TIME: 14:23:13

Input Set : A:\797,553SeqList.txt

Output Set: N:\CRF4\10182004\J797553A.raw

3 <110> APPLICANT: Moyle, William R.  
 4 Xing, Yonga  
 6 <120> TITLE OF INVENTION: Protein Knobs  
 8 <130> FILE REFERENCE: Moy1-005US  
 10 <140> CURRENT APPLICATION NUMBER: 10/797,553A  
 11 <141> CURRENT FILING DATE: 2004-03-10  
 13 <150> PRIOR APPLICATION NUMBER: PCT/US02/35914  
 14 <151> PRIOR FILING DATE: 2002-11-08  
 16 <160> NUMBER OF SEQ ID NOS: 66  
 18 <170> SOFTWARE: PatentIn version 3.3  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 92  
 22 <212> TYPE: PRT  
 23 <213> ORGANISM: Artificial  
 25 <220> FEATURE:  
 26 <223> OTHER INFORMATION: Artificial Sequence  
 28 <400> SEQUENCE: 1  
 30 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
 31 1 5 10 15  
 34 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
 35 20 25 30  
 38 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
 39 35 40 45  
 42 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
 43 50 55 60  
 46 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
 47 65 70 75 80  
 50 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
 51 85 90  
 54 <210> SEQ ID NO: 2  
 55 <211> LENGTH: 92  
 56 <212> TYPE: PRT  
 57 <213> ORGANISM: Artificial  
 59 <220> FEATURE:  
 60 <223> OTHER INFORMATION: Artificial Sequence  
 62 <400> SEQUENCE: 2  
 64 Ala Pro Asp Val Cys Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
 65 1 5 10 15  
 68 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
 69 20 25 30  
 72 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
 73 35 40 45  
 76 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser

(ps.6)

Does Not Comply  
Corrected Diskette NeededPLS explain source  
OF genetic material.

Invalid

Response

PLS explain source of genetic  
material.

Invalid

Response

↑ PLS see item #11 on  
 error summary sheet. 10/18/04

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/797,553A

TIME: 14:23:13

Input Set : A:\797,553SeqList.txt

Output Set: N:\CRF4\10182004\J797553A.raw

77      50                      55                      60  
80 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
81 65                      70                      75                      80  
84 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
85                      85                      90  
88 <210> SEQ ID NO: 3  
89 <211> LENGTH: 92  
90 <212> TYPE: PRT  
91 <213> ORGANISM: Artificial  
93 <220> FEATURE:  
94 <223> OTHER INFORMATION: artificial sequence  
96 <400> SEQUENCE: 3  
98 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Cys Gln Glu Asn Pro  
99 1                      5                      10                      15  
102 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
103                      20                      25                      30  
106 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
107                      35                      40                      45  
110 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
111                      50                      55                      60  
114 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
115 65                      70                      75                      80  
118 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
119                      85                      90  
122 <210> SEQ ID NO: 4  
123 <211> LENGTH: 92  
124 <212> TYPE: PRT  
125 <213> ORGANISM: Artificial  
127 <220> FEATURE:  
128 <223> OTHER INFORMATION: artificial sequence  
130 <400> SEQUENCE: 4  
132 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Cys Pro  
133 1                      5                      10                      15  
136 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
137                      20                      25                      30  
140 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
141                      35                      40                      45  
144 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
145                      50                      55                      60  
148 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
149 65                      70                      75                      80  
152 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
153                      85                      90  
156 <210> SEQ ID NO: 5  
157 <211> LENGTH: 92  
158 <212> TYPE: PRT  
159 <213> ORGANISM: Artificial  
161 <220> FEATURE:  
162 <223> OTHER INFORMATION: artificial sequence

✓ SAME  
errors

✓ PLS  
see  
item #  
= on error  
summary  
sheet.

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/797,553A

TIME: 14:23:13

Input Set : A:\797,553SeqList.txt

Output Set: N:\CRF4\10182004\J797553A.raw

*same errors*

164 <400> SEQUENCE: 5  
166 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
167 1 5 10 15  
170 Cys Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
171 20 25 30  
174 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
175 35 40 45  
178 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
179 50 55 60  
182 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
183 65 70 75 80  
186 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
187 85 90  
190 <210> SEQ ID NO: 6  
191 <211> LENGTH: 92  
192 <212> TYPE: PRT  
193 <213> ORGANISM: Artificial  
195 <220> FEATURE:  
196 <223> OTHER INFORMATION: artificial sequence  
198 <400> SEQUENCE: 6  
200 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
201 1 5 10 15  
204 Phe Phe Ser Gln Pro Cys Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
205 20 25 30  
208 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
209 35 40 45  
212 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
213 50 55 60  
216 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
217 65 70 75 80  
220 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
221 85 90  
224 <210> SEQ ID NO: 7  
225 <211> LENGTH: 92  
226 <212> TYPE: PRT  
227 <213> ORGANISM: Artificial  
229 <220> FEATURE:  
230 <223> OTHER INFORMATION: artificial sequence  
232 <400> SEQUENCE: 7  
234 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
235 1 5 10 15  
238 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Cys Cys Met Gly Cys Cys  
239 20 25 30  
242 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
243 35 40 45  
246 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
247 50 55 60  
250 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
251 65 70 75 80

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/797,553A

DATE: 10/18/2004

TIME: 14:23:13

Input Set : A:\797,553SeqList.txt

Output Set: N:\CRF4\10182004\J797553A.raw

*VSAMP  
errors*

254 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
 255 85 90  
 258 <210> SEQ ID NO: 8  
 259 <211> LENGTH: 92  
 260 <212> TYPE: PRT  
 261 <213> ORGANISM: Artificial  
 263 <220> FEATURE:  
 264 <223> OTHER INFORMATION: artificial sequence  
 266 <400> SEQUENCE: 8  
 268 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Cys Gln Glu Asn Pro  
 269 1 5 10 15  
 272 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
 273 20 25 30  
 276 Phe Ser Arg Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
 277 35 40 45  
 280 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
 281 50 55 60  
 284 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
 285 65 70 75 80  
 288 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
 289 85 90  
 292 <210> SEQ ID NO: 9  
 293 <211> LENGTH: 92  
 294 <212> TYPE: PRT  
 295 <213> ORGANISM: Artificial  
 297 <220> FEATURE:  
 298 <223> OTHER INFORMATION: artificial sequence  
 300 <400> SEQUENCE: 9  
 302 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
 303 1 5 10 15  
 306 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys  
 307 20 25 30  
 310 Phe Ser Cys Ala Tyr Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu  
 311 35 40 45  
 314 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser  
 315 50 55 60  
 318 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr  
 319 65 70 75 80  
 322 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser  
 323 85 90  
 326 <210> SEQ ID NO: 10  
 327 <211> LENGTH: 92  
 328 <212> TYPE: PRT  
 329 <213> ORGANISM: Artificial  
 331 <220> FEATURE:  
 332 <223> OTHER INFORMATION: artificial sequence  
 334 <400> SEQUENCE: 10  
 336 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro  
 337 1 5 10 15

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/797,553A

TIME: 14:23:13

Input Set : A:\797,553SeqList.txt

Output Set: N:\CRF4\10182004\J797553A.raw

*CSAME  
errors*

```

340 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys
341           20           25           30
344 Phe Ser Arg Ala Cys Pro Thr Pro Leu Arg Ser Lys Lys Thr Met Leu
345           35           40           45
348 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser
349           50           55           60
352 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr
353 65           70           75           80
356 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser
357           85           90
360 <210> SEQ ID NO: 11
361 <211> LENGTH: 92
362 <212> TYPE: PRT
363 <213> ORGANISM: Artificial
365 <220> FEATURE:
366 <223> OTHER INFORMATION: artificial sequence
368 <400> SEQUENCE: 11
370 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro
371 1           5           10           15
374 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys
375           20           25           30
378 Phe Ser Arg Ala Tyr Cys Thr Pro Leu Arg Ser Lys Lys Thr Met Leu
379           35           40           45
382 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser
383 50           55           60
386 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr
387 65           70           75           80
390 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser
391           85           90
394 <210> SEQ ID NO: 12
395 <211> LENGTH: 92
396 <212> TYPE: PRT
397 <213> ORGANISM: Artificial
399 <220> FEATURE:
400 <223> OTHER INFORMATION: artificial sequence
402 <400> SEQUENCE: 12
404 Ala Pro Asp Val Gln Asp Cys Pro Glu Cys Thr Leu Gln Glu Asn Pro
405 1           5           10           15
408 Phe Phe Ser Gln Pro Gly Ala Pro Ile Leu Gln Cys Met Gly Cys Cys
409           20           25           30
412 Phe Ser Arg Ala Tyr Pro Cys Pro Leu Arg Ser Lys Lys Thr Met Leu
413           35           40           45
416 Val Gln Lys Asn Val Thr Ser Glu Ser Thr Cys Cys Val Ala Lys Ser
417 50           55           60
420 Tyr Asn Arg Val Thr Val Met Gly Gly Phe Lys Val Glu Asn His Thr
421 65           70           75           80
424 Ala Cys His Cys Ser Thr Cys Tyr Tyr His Lys Ser
425           85           90
428 <210> SEQ ID NO: 13

```

The type of errors shown exist throughout  
the Sequence Listing. Please check subsequent  
sequences for similar errors.

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/797,553A

DATE: 10/18/2004  
TIME: 14:23:14

Input Set : A:\797,553SeqList.txt  
Output Set: N:\CRF4\10182004\J797553A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:56; Xaa Pos. 1,8,10

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27

Seq#:28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51

Seq#:52,53,54,55,56,57,58,59,60,61,62,63,64,65,66



**VERIFICATION SUMMARY**

DATE: 10/18/2004

PATENT APPLICATION: US/10/797,553A

TIME: 14:23:14

Input Set : A:\797,553SeqList.txt

Output Set: N:\CRF4\10182004\J797553A.raw

L:2208 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0